BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 6: e93803
doi: 10.3897/biss.6.93803

OPEN ACCESS

Conference Abstract

Using Crowd-sourcing Platforms to Increase and Spread Knowledge on the Biodiversity in Sub-Saharan Africa

Mohammed Kamal Deen Fuseini Dnshitobu‡, Agnes Ajuma Abah§, Benedict Udeh§, Andra Waagmeester|

‡ Wiki Mentor Africa, Tamale, Ghana
§ Wiki Mentor Africa, Abuja, Nigeria
| Wiki Mentor Africa, Ekeren, Belgium

Corresponding author: Mohammed Kamal Deen Fuseini Dnshitobu (fuseinikamaldeen@gmail.com), Andra Waagmeester (andra@micelio.be)

Received: 22 Aug 2022 | Published: 23 Aug 2022

Citation: Dnshitobu MKDF, Abah AA, Udeh B, Waagmeester A (2022) Using Crowd-sourcing Platforms to Increase and Spread Knowledge on the Biodiversity in Sub-Saharan Africa. Biodiversity Information Science and Standards 6: e93803. https://doi.org/10.3897/biss.6.93803

Abstract

Sub-Saharan Africa possesses a wide range of natural habitats and climatic regions. Understanding the biodiversity in sub-Saharan Africa is still at an early stage. Action is therefore needed at all levels of biodiversity science to write, educate and understand the benefits of improved knowledge. While there is already substantial knowledge on biodiversity globally, the knowledge is often concentrated in specific areas in our world. Fig. 1 is a map of occurrences of flora and fauna recorded in the Global Biodiversity Information Facility (GBIF). It shows that the recordings are unevenly distributed. Occurrences from Ghana and Nigeria seem to be underrepresented.

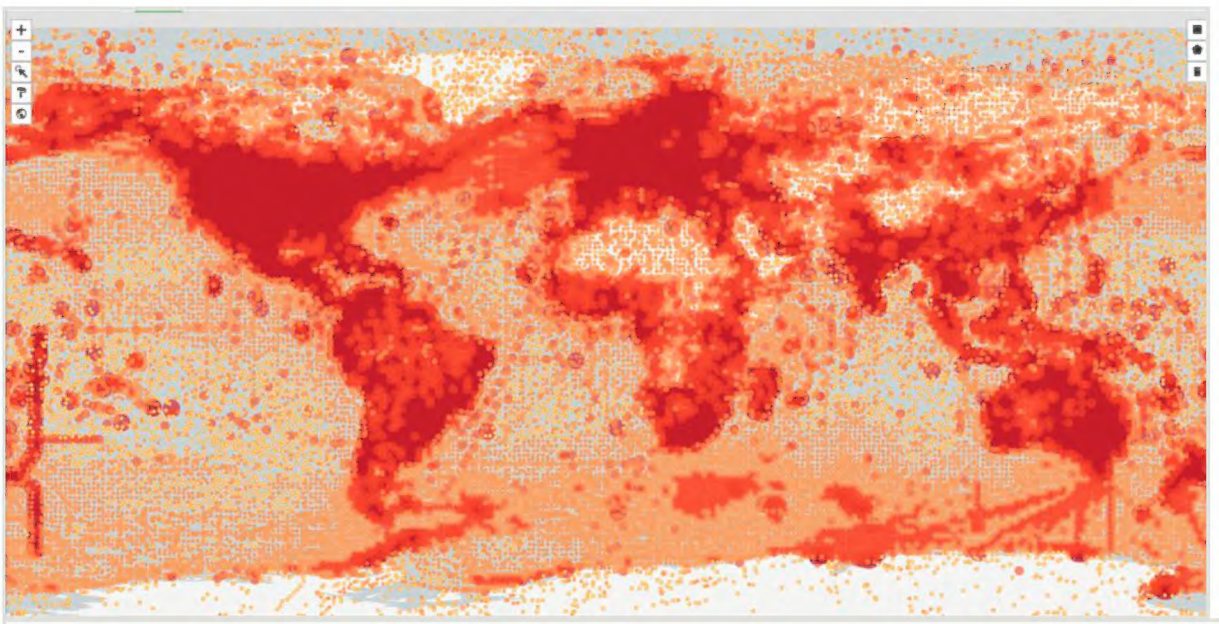

Figure 1.

Observations in GBIF (seen: June 30, 2022).

Access to available knowledge on biodiversity can also be limited. Wikipedia is a popular source of knowledge and it is also available in different languages, which can help in spreading the knowledge across languages. These different language editions are linked through Wikidata, the linked data repository associated with Wikipedia. Wikidata currently has almost 3,5 million records on taxa. Many species lack a Wikipedia article. This number of missing Wikipedia articles is low when looking at the number of articles in local languages of sub-Saharan Africa on July 1, 2022 (Table 1).

Table 1.

Number of Wikipedia articles on taxa in a select set of sub-Saharan languages (sample Wikidata query: https://w.wiki/5NpX).

Language	Wikipedia	articles
Swahili	https://sw.wikipedia.org/	903
Hausa	https://ha.wikipedia.org/	444
Bambara	https://bm.wikipedia.org/	431
Amharic	https://am.wikipedia.org/	415
Dagbani	https://dag.wikipedia.org/	392
Yoruba	https://yo.wikipedia.org/	366
Malagasy	https://mg.wikipedia.org/	287
Ewe	https://ee.wikipedia.org/	255
Sango	https://sg.wikipedia.org/	237
Dinka	https://din.wikipedia.org/	149
Zulu	https://zu.wikipedia.org/	124
Wolof	https://wo.wikipedia.org/	120
Kirundi	https://rn.wikipedia.org/	78

Language	Wikipedia	articles
Shilha	https://shi.wikipedia.org/	73
Kongo	https://kg.wikipedia.org/	71
Chewa	https://ny.wikipedia.org/	56
Kabiye	https://kbp.wikipedia.org/	48
Tsonga	https://ts.wikipedia.org/	38
Tumbuka	https://tum.wikipedia.org/	28
Igbo	https://ig.wikipedia.org/	18
Xhosa	https://xh.wikipedia.org/	16
Gikuyu	https://ki.wikipedia.org/	13
Sesotho	https://st.wikipedia.org/	9
Akan	https://ak.wikipedia.org/	8
Venda	https://ve.wikipedia.org/	8
Swazi	https://ss.wikipedia.org/	7
Twi	https://tw.wikipedia.org/	6
Tswana	https://tn.wikipedia.org/	4
Gun	https://guw.wikipedia.org/	1

We can conclude that there are significant gaps in our knowledge about taxonomy of Sub-Saharan species, as a significant number of species remain undescribed. Even among described species, there is a distinct lack of knowledge regarding their distribution and biogeography, as well as basic biology, such as life histories, feeding habits and habitat preferences. Existing knowledge is spread across different data silos and across different languages, which makes it challenging to get a complete overview of the existing knowledge on local biodiversity. We present a sub-project in the community project Wiki Mentor Africa. Here we record observations of different species and spread the acquired knowledge in Dagbani and Igbo, two languages spoken in the sub-Saharan region. For this we use two crowd-sourcing platforms, iNaturalist and the Wikimedia Stack comprising Wikipedia, Wikimedia Commons and Wikidata. iNaturalist is a community-driven, free-access platform for sharing observations of the surrounding flora and fauna, accessible through a dedicated website and mobile applications (Android & iOS). On July 1st 2022 the iNaturalist community contained more than 106 million observations and more than 2.2 million users. Observations can be identified by automatic suggestion or confirmation by fellow observers. When identified observations are confirmed, they are tagged as being research-grade. While caution is advised, research-grade observations in iNaturalist can act as illustrations in Wikipedia articles on the associated flora and fauna. While the default license does not allow this reuse, users can change their settings to a compatible license. We are developing workflows that search iNaturalist for compatibly licensed photos (e.g., Creative Commons licenses CC BY, CC BY-SA, and public domain licenses such as CC0) for reuse in Wikipedia. This is possible since there are active links maintained between iNaturalist, Wikidata and Wikipedia.

In Wikidata, taxa and their respective iNaturalist identifiers are mapped. This allows identification of missing Wikipedia articles in any of the approximately 300 supported languages in Wikipedia.

We have developed a Jupyter notebook that uses the active links between the different platforms to identify missing Wikipedia articles. This notebook collects data from different biodiversity-related databases, such as GBIF or the Biodiversity Heritage Library (BHL), to prepare an initial stub for the Wikipedia article to be created. These so-called stubs can then be extended by those respective language communities.

Keywords

Wikidata, iNaturalist, multilingualism, Wikipedia

Presenting author

Mohammed Kamal-Deen Fuseini Dnshitobu

Presented at

TDWG 2022

Funding program

Wiki Mentor Africa https://www.wikidata.org/wiki/Wikidata:Wiki_Mentor_Africa